

Strained Mercy: The Quality of Medical Care in Delhi

Jishnu Das

Development Research Group, World Bank

jdasl@worldbank.org

Jeffrey Hammer

Development Research Group, World Bank

jhammer@worldbank.org

with

The Institute of Socio-Economic Research on Development and Democracy

Delhi

World Bank Policy Research Working Paper 3228, March 2004

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent. Policy Research Working Papers are available online at <http://econ.worldbank.org>.

The modules used in this study were designed in consultation with Dr. Tejvir Singh Bubba Khurana and many discussions with Ken Leonard and Asim Khwaja. The pilot and survey was implemented by Jishnu Das and Jeffrey Hammer with N. Deepak, Pritha Dasgupta, Sourabh Priyadarshi, Poonam Kumari and Sarasij Majumdar, all members of The Institute of Socio-Economic Research on Development and Democracy Delhi (ISERDD). Further support from Purshottam, Rajan Kumar and Simi Bajaj, often under trying circumstances, is gratefully acknowledged. We also thank Dr. Arvind Taneja, Veena Das, and R. K. Das for comments and suggestions; to the panel of physicians led by Dr. Jonathan Ellen and Dr. Zahida Khwaja for their cooperation in evaluating treatments; and to Shruti Haldea for excellent research assistance. Finally, the project would not have been possible without the cooperation and enthusiasm of the participating providers as well as administrators of the various public sector facilities surveyed. The research was funded by a research grant from the World Bank.

Abstract

The quality of medical care is a potentially important determinant of health outcomes. Nevertheless, it remains an understudied area. The limited research that exists defines quality either on the basis of drug availability or facility characteristics, but little is known about how *provider* quality affects the provision of health care. We address this gap through a survey in Delhi with two related components. We evaluate “competence” (what providers know) through vignettes and practice (what providers do) through direct clinical observation. Overall quality, as measured by the competence necessary to recognize and handle common and dangerous conditions, is quite low albeit with tremendous variation. While there is some correlation with simple observed characteristics, there is still an enormous amount of variation within such categories. Further, even when providers know what to do they often don’t do it in practice. This appears to be true in both the public and private sectors but for very different, and systematic, reasons. In the public sector providers are more likely to commit errors of omission—exert less effort compared to their private counterparts. In the private sector providers are prone to errors of commission—they are more likely to behave according to the patient’s expectations resulting in the inappropriate use of medications, the overuse of antibiotics, and increased expenditures. This has important policy implications for our understanding of how market failures and failures of regulation in the health sector affect the poor.

Contents

I.	Introduction	1
II.	What Does Quality Mean? Understanding Competence and Practice	5
III.	Measuring Quality: The Method and Implications	7
III.1.	Sampling	7
III.2.	Assessing Competence: Methodology	8
III.3.	Assessing Competence: Benchmarking	11
III.4.	What Does the Distribution of Competence Imply for Care?	12
IV.	What Providers Do: The Relationship between Competence and Practice?	15
IV.1.	Competence and Practice: Is There a Difference?	16
IV.2.	Incentives and the “Wedge”: Evidence from Public/Private Providers	19
V.	Discussion and Conclusion	24
	Bibliography	26

I. Introduction

For all its acknowledged importance, the quality of medical care in India is not well understood. Many studies of the demand for medical services identify quality as a major determinant of facility selection and use, and that such use should contribute to improvements in health status. However, quality of care is not clearly defined beyond easily observable facility characteristics such as the availability of medicines or functioning equipment (Collier, Dercon, and Mackinnon 2003; Lavy and Germain 1994).¹ What is less well understood, but has greater relevance for both the use and outcomes of services, is the quality of the actual advice given in clinical settings. Medicine can be bought over the counter and is not specific to particular facilities or providers. But good advice from a provider with knowledge and experience is a service that patients cannot easily obtain by themselves.²

While clearly important, it is understandable why the clinical quality of medical advice is not well studied. It is not easy to define what should be measured and even with a clear definition, measurement is hard to implement in practice. Our paper addresses this gap and presents results of research on different dimensions of quality as applied to medical caregivers in Delhi. We distinguish between quality measured as “What do medical care providers know?” and as “What do medical care providers do?” We refer to the first as “competence” and the second as “practice.” The wedge between competence and practice—the relation between what they know and what they do—then raises important theoretical and empirical questions related to the structure of the health care market.³

¹ This observation is restricted to research in economics—there is a large literature on the quality of prescriptions and medication in India, starting with Greenhalgh (1987) and with more recent contributions by Kamat and Nichter (1998) and Phadke (1998). These studies, which document the patterns of prescription and self-medication in the Indian health environment, suffer from two related problems. First (with the notable exception of Phadke 1998) there is little information linking the prescription to the characteristics of the provider. Second (partly as a result of this omission) there is little that can be said from these studies on how prescription practices can be improved, i.e., does the dramatic overuse of antibiotics stem from patient demand, overall uncertainty in the health environment, or other factors such as doctor/patient ignorance? Understanding the mechanisms that give rise to observed patterns of medication is a necessary first step in framing policy in the health sector.

² Theoretically this refers to the difference between goods that can be traded and those that cannot. If medicine is not available in the hospital, individuals can buy it from the pharmacist. Human capital on the other hand is harder to trade—advice cannot be purchased without the presence of the doctor. Thus while free medication will lead to greater use of the facility (equivalent to a pharmacist who “sells” his medicines free of cost), the impact on health *outcomes* depends entirely on the quality of care imparted.

³ This article summarizes findings in two recent papers: Das and Hammer (2003a, and 2003b)

To answer these questions, we analyze two types of data collected from more than 200 medical care providers in seven neighborhoods of the city.⁴ The first, on provider competence, is based on a set of “vignettes,” or five hypothetical cases presented to providers who were asked what they would do in each circumstance—what questions they would ask, what examinations they would perform, and what treatments they would recommend. The second, on provider practice, was based on direct observation of their interactions with patients. Interviewers sat with providers during their clinical session for a whole day noting the number of patients, their complaints, their questions, examinations performed, and the treatments (including specific drug names). In both the public and the private sector, prices charged were recorded as well.⁵ We then studied the relationship between the two sets of answers to see whether they were consistent and, if not, what other factors could explain performance.

The analysis was motivated by several concerns. The first was to see if there were major differences in competence and practice along several dimensions of provider characteristics such as their training and qualifications, their location in poorer or more affluent neighborhoods and whether they were in the public or private sector. The second concern, related to the problem of public and private sectors, was whether practice deviated from competence according to the different incentives for performance—doctors on salary have different motivations than those receiving a fee for service. The third concern was about the general ability to diagnose and deal with a spectrum of diseases ranging from common problems such as diarrhea to life-threatening conditions such as tuberculosis. Finally, we were interested in the over-use of medications, particularly of antibiotics, since it is a widely acknowledged problem in the country.

The results are depressing in at least two ways. First, overall quality as measured by the competence necessary to recognize and handle common and dangerous conditions is quite low, although it varies from very good to unimaginably low. Further, while there is some correlation with simple observed characteristics—the mean level of competence in providers with a degree in biomedicine (MBBS) or providers located in high-income neighborhoods is higher than those with other qualifications and/or located in poorer neighborhoods—there is still considerable variation within each category. Thus, it is difficult to predict the level of provider competence based solely on qualifications and location.

⁴ A medical care provider is defined as any individual or institution that provides medical care on a professional basis. We use the phrase “provider” instead of “doctor” since our sample includes both trained and untrained individuals.

⁵ Payments in public facilities were recorded but virtually none were observed—whether this reflects actual practice or the fact that we were watching is unknown.

Second, even when providers know in theory what treatment protocol to follow, they frequently deviate from the protocol in practice. This startling observation holds for both the public and private sectors but for very different, and systematic, reasons. While neither provider spends much time with their patients—averaging about three minutes per contact—public providers spend much less than private counterparts. They ask fewer questions, do fewer examinations, and provide less treatment. Private providers do more of each. This contrast is all the more striking in that we find that patients presenting in the public sector have, on average, a longer history of illness than those in the private sector and (from a parallel household survey) this is often because they have already consulted the private sector but have failed to report improvements in their respective conditions.

Given that private providers exert more effort than their public counterparts, can one conclude that private providers are better than public? The brief answer is “not necessarily.” From the patient’s point of view there is a price differential that could compensate for the difference in care received. A more subtle argument stems from the recognition that in medicine, both “less is more” and “more is more” can be true, depending on the nature of the ailment. That is, there are two types of errors. Type I errors arise when providers do something when it is not called for (as in over-medication) and Type II errors arise when providers do not do something that *is* called for. The losses from Type I errors can be potentially large. In most disease environments, close to 80 percent of all illnesses are “self-limiting” in the sense that no medication (apart from those given for symptomatic relief) is required for a cure. Type I errors then imply large expenditures on minor ailments (with particularly dire economic consequences for the poor) and in the cases where such treatments include antibiotics or steroids, a worsening of health outcomes due to microbial resistance in the long run. We find the private sector particularly prone to Type I errors and the public sector prone to Type II errors.

To elaborate on this point further, consider the following excerpt from an interview (Das) with a popular provider in West Delhi.

“Dr. S’s clinic is opposite a main road in West Delhi. The board outside advertises the services of Dr. P, a gold medal winner in the MBBS exams, but every time we have gone, we have found only Dr. S and his wife, sitting on either side of the table with an array of bottles containing tablets and syrups around them, and a mortar and pestle on the table. The clinic is always extremely busy—probably the busiest in the locality, with Dr. S estimating that he and his wife together see between 200 and 400 patients a day. Patients sit on a bench that spills into the street, and as each one is given medicine the entire group shifts forward one to make space for the latest entrant. The patient that is currently being seen gets a privileged space at the front of the bench, separating him/her by a foot or so of space from the next candidate.

Dr. S.: “What we are doing here is charity. I never charge more than Rs. 2 or so from these patients and give them the medicines free of charge.”

DAS: “We have seen a lot of diarrhea cases among our households these days.”

Dr. S.: “Yes, there is a lot of diarrhea and dysentery in this locality—what can they do as well? The water is dirty and people do not know to boil it—that’s why their children are always falling sick.”

DAS: “So, what do you do for children with diarrhea?”

Dr. S.: “What can we do? The usual things—we tell the mother to give water with salt and sugar to the baby and then also give some medicines.”

DAS: “Such as?”

Dr. S.: “The usual—metrogyl (metronidazole), loperamide (an anticholinergic), Furoxone (furazolidone).”

DAS: “But isn’t ORS enough?”

Dr. S.: “Of course the WHO and others keep saying that we should only give ORS. But if I tell the mother that she should go home and only give the child water with salt and sugar, she will never come back to me; she will only go to the next doctor who will give her all the medicines and then she will think that he is better than me.”

In the private sector providers feel constrained in their behavior by the nature of the market, so that even if they *know* the correct course of action, they may choose to act differently to cater to the patient’s expectations. This is troubling since economic reasoning would lead us to expect that the highly competitive nature of the market (we found at least 70 providers within a 15-minute walk of any of the seven localities in our study) would result in “efficient” outcomes so that high quality care is provided at the lowest cost. The interview with Dr. S. as well as our data on provider-patient interactions indicate that the private market encourages incorrect treatments *even among providers who are highly competent*. In such an environment the relative cost of Type I and Type II errors complicates substantially the evaluation of the private versus the public sector.

Policy options in such an environment are particularly difficult to determine. For the patient who does not have any means of assessing the actual level of provider competence, “average” competence comes to be inferred from the general notion about the private sector or public hospitals (for instance, the relationship between competence and qualification or locality). Yet the variance in the level of competence combined with the difference between competence and practice implies that the actual efficacy of treatment is determined more by the type of illness a patient has than her ability to choose the most competent provider to treat it.

Thus patients with either a self-limiting illness or one that requires a referral to specialist care will generally be served better in the public sector because providers in this sector treat with fewer medicines and are likely to refer the patient to another facility in case of severe or complicated cases. If active, primary-level care is called for,

however, the private sector is often the better option since the patient is likely to receive some treatment for her symptoms. However, people rarely know the exact nature of their illness before seeking advice and are hence not in a position to make an informed choice even if they were able to decipher these particular characteristics of the private versus public facilities in health care. The practical difficulties of a systematic regulatory framework for medical care are very severe. This leaves an important role for improving the ability of patients to make better choices among their options, especially since there is no dearth of potential providers to choose from in the urban Indian context. However, the problem of how to convey the appropriate information to people is daunting.

The remainder of our paper is structured as follows. Section II briefly discusses the potential sources of differences in competence and practice. Section III details the method and basic results from the measurement of competence and section IV studies the relationship between provider practice and competence. Section V concludes with a brief discussion of policy options.

II. What Does Quality Mean? Understanding Competence and Practice

Consider a provider faced with a patient with a set of symptoms, but an undiagnosed (as yet) illness. The outcome of their interaction (is the patient cured?) will clearly depend on the competence of the provider and the illness of the patient (referred to in the literature as the *case mix*), but will probably be affected by other factors as well. Two examples are *patient compliance* and *provider incentives*.

Compliance refers to the idea that patients may not fully comply with the provider's instructions because the cost of doing so (possibly in terms of effort) are high. As an example, depending on patient characteristics, many physicians in the United States prefer to treat patients with sexually transmitted diseases with a single-use antibiotic rather than with antibiotics that require a longer course. The former strategy is optimal when there is a high likelihood that the patient will not comply or return for future consultations.

Similarly, provider incentives relate to the idea that the current treatment of the patient affects future prospects of the provider. As an example, a physician in the public sector who has no incentive to induce the patient to return will behave differently from one in the private sector, whose future profits are directly linked to whether the patient returns for treatment (either for this or another illness) in the future.

It is important to underscore that case mix affects the relationship between competence and health outcomes. If patients with more severe diseases choose to consult more competent providers, then case outcomes will not provide a good measure of

competence since competent providers are receiving a disproportionately higher number of patients with severe diseases. If selection alone were the problem, standardizing the case mix by observing the same illness in both the measurement of competence and in clinical observation, should eliminate any difference between competence and practice.

Recent literature, however, provides fairly strong evidence that other factors matter as well. Rethans and others (1991) report significant differences between competence and practice in a study where doctors were assessed through standardized patients followed by vignettes, although the authors are unable to attribute this to any one specific characteristic of the provider. Similarly, Leonard and Masatu (2003) show that physician contracts in Africa are designed to address problems arising from effort costs on the providers side and compliance costs on the patients: illnesses that require greater patient compliance are associated with contracts characterized by higher up-front payments and in-patient care. In a separate study, Leonard (2003) also documents that competence is distinct from practice, with physicians undertaking fewer physical examinations in a clinical setting compared to the same case administered through vignettes.

Our study proceeds in a similar vein. We first administered vignettes to providers in our sample in order to measure competence. These vignettes are standardized, hypothetical cases of specific diseases with some basic facts about the case. We then record what the providers ask about the history of the client and the illness, what kinds of examinations they call for (with answers anticipated and given at once), and what treatment is recommended. Further, vignettes also allowed us to standardize patient characteristics *in addition* to the case mix so that a measure of competence uncontaminated by other patient or provider characteristics was obtained. For instance, to control for compliance behavior the provider was informed that the patient would fully comply with all recommendations and return to the provider if required. Second, to measure practice we use direct clinical observation whereby some basic facts about each interaction between the provider and (real) patients are recorded.

These multiple sources of data on competence and practice can be used in a number of ways. As a first step we can evaluate the impact of competence on provider behavior in the clinic. Do more competent providers behave in systematically different ways, either in terms of the time spent with the patient or the likelihood of physical examinations? Next, we can examine the wedge between competence and practice. Using the questions and the treatments provided in the vignettes, we can compare responses obtained to what is undertaken in practice. While documentation alone of this difference is important, the implications either for care-seeking or health care policy remain unclear—does this wedge represent an “optimal” tradeoff between cure probabilities and the cost of effort or are their market inefficiencies that cause this difference to arise? To

address this issue, we analyze how this wedge differs among different providers in our sample.

In particular, if incentives affect the behavior of providers, we expect to find that private providers are more sensitive to the patient's wants (and hence getting the patient to return) than their public counterparts. To the extent that this implies a higher standard of care, the welfare of the patient is enhanced. However, to the extent that this leads to the systematic over-provision of medication, or inappropriate medication, welfare is reduced. Assessing the relative impact of each in relation to provider competence becomes critical for developing public-private partnerships in health care.

III. Measuring Quality: The Method and Implications

We present our results in four sections. We start by briefly discussing the sampling and statistical methodology used to measure competence. Second, we benchmark the measure of quality obtained with treatment patterns to assess what quality *means* in real terms such as “a provider with a high rank in our sample will detect tuberculosis x percent of the time.” We then turn to the distribution of competence across neighborhoods and income categories with a focus on the difference between public and private providers. Finally, we discuss how our competence ranks correlate with what providers do in practice. This wedge is analyzed in some detail with reference to incentives in the public and private sectors.

III.1. Sampling

The sample of providers interviewed was drawn in parallel with a longitudinal household study conducted by the Institute of Socio-Economic Research on Democracy and Development (ISERDD) in Delhi. This study followed 300 families in seven neighborhoods in Delhi with weekly visits for a period of 18 months, asking about their health problems and their provider and treatment choices. This generated a very long list of providers visited. In addition to this list, we undertook a census of providers—research staff systematically completed a basic questionnaire of *all medical providers* within a radius of a 15-minute walk from the edges of the neighborhood. For all providers, either visited by households in the sample or located within a 15-minute radius, a short interview was administered to check qualifications (degrees, if any, length of time in the community) and they were then part of the universe of providers chosen for more detailed interviews. A total of 571 providers formed the list frame for our sampling.

From this overall list, we randomly chose 25 providers in each of the seven neighborhoods from the first group (those visited by the households) with weights proportional to the number of visits. In addition, 10 providers who had never been visited

by any of the households but were present in the provider census in each locality were randomly chosen for a total sample of 235. Vignettes and clinical observations were completed for 205; the remainder had either left the locality, were specialists (such as ophthalmologists), or did not give permission for the survey. Since we were interested in the average provider found in different localities of Delhi, our sample is not restricted only to MBBS doctors, but includes providers in other streams such as Ayurveda (BAMS or BIMS) and Unani (BUMS) practitioners as well as those who are registered practitioners but have no formal medical training (RMP or PMP).⁶

III.2. Assessing Competence: Methodology

Our measure of clinical competence combined techniques from two different disciplines: vignettes from the medical literature and scoring techniques from the literature on educational testing.⁷ To administer vignettes we sent a pair of interviewers for each interview—one to act as the patient (or the patient’s mother in the case of diarrhea) and one to record responses and to answer questions posed by the doctor that a patient would not be in a position to know such as, for example, the results of blood tests.

Each interview involved five vignettes: (1) a case of simple diarrhea for an infant; (2) a case of viral pharyngitis for an adult; (3) a case of tuberculosis; (4) a case of depression and (5) a case of pre-eclampsia (a life threatening condition of high blood pressure in a pregnant woman). Case I is included since it is a very common problem where possible serious causes must be ruled out, but treatment is limited to oral rehydration therapy (ORT). Case II is included to check for tendencies for over-medication. Case III checks for the ability to treat or refer a life-threatening infectiousness condition. Case IV also checks for tendency to over-medicate and ability to diagnose non-medical problems. Case V is a life-threatening complication of pregnancy and must be identified and referred to a hospital immediately.

For each of the cases the “patient” would start off by describing the illness. Thus in the case of diarrhea the doctor is informed that *“A mother brings in an 8-month old male child to you. The mother complies with all tests and medications that you recommend and will return to you if you require.”* The “mother” then says *“My child has been suffering from diarrhea for the last two days, and I do not know what to do.”* From this point on, the provider is invited to proceed exactly as he/she would with a normal patient. She may thus ask questions regarding the history of the illness as well as potential examinations (blood tests) and for each question asked, a standardized answer is provided.

⁶ The interested reader is referred to Das and Das (2003) for a more detailed description on the types of providers found in the localities as well as the genesis of different qualifications in Indian health policy.

⁷ Details can be found in Das and Hammer (2003a).

When evaluating provider scores, the entire interaction is taken into account—what history questions were asked, what examinations were performed, and what treatment was given. “Appropriate” treatment was judged by three panels of physicians, two from South Asia and one from Johns Hopkins University in the United States. In each case the team of highly trained physicians who evaluated the treatments acted independently, so that consistency of the desired treatment protocol, i.e., the extent to which physicians in different epidemiological environments agreed on a standardized treatment protocol, could later be assessed through cross-rater agreements. Due to the straightforward nature of the cases, there was virtually no disagreement as to the proper course of treatment. Table 1 presents examples from these cases.

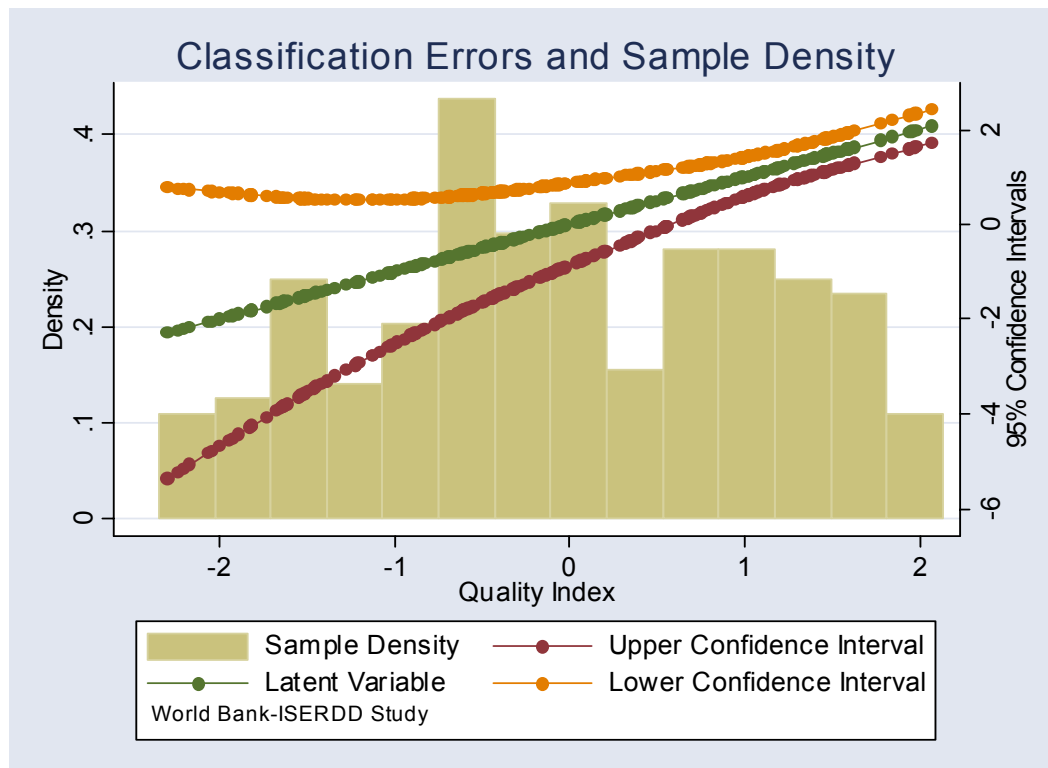
The second element of our methodology subjected the responses in the vignettes to Item Response Theory (IRT), commonly used in the literature on educational attainment. This method extracts from a battery of questions an overall measure of an underlying variable (degree of knowledge) for each of the respondents. We ended up using 79 questions—those asked frequently enough to stand up to the statistical analysis—over the

Table 1: Examples of Vignettes Used			
Case	History question examples (doctor is evaluated positively for asking)	Examination/tests examples (doctor is evaluated positively for)	Recommended treatment (doctor is evaluated positively for)
I. Simple diarrhea in an infant	Whether there is blood or mucous in the stool (indications of possible bacterial infections).	Checking for signs of dehydrations (either through skin-tenting or depression of the skull fontanel).	Oral Rehydration Therapy and no anti-infective or anti-diarrheal medicines.
II. Viral Pharyngitis in adult man	Color of sputum, respiration rate (to rule out bacterial infections)	Check for fever, breathing difficulty.	Recommending rest without further treatment, although symptomatic relief for cough was not penalized by the expert panels.
III. Tuberculosis in adult man	Color of sputum (flecks of red indicate potential TB)	Sputum test and/or chest x-ray (results are positive)	Proper drug therapy or referral to public TB clinic
IV. Depression in young woman	Thoughts of suicide? (anti-depressants are not immediately effective)	Examination of thyroid gland (to rule out alternative explanations)	Referral for therapy or counseling (but no rush)
V. Pre-eclampsia (pregnant woman)	Duration of headache, duration of pregnancy	Blood pressure (result is very high)	Immediate referral to hospital
<i>Note:</i> This table provides examples of questions that the provider had to ask and treatments that she had to provide for a positive evaluation. Weights assigned to each question were assigned through the IRT methodology as described in the test and in greater detail in Das and Hammer (2003a).			

five cases. This method is related to other statistical techniques such as principal components analysis or factor analysis; an important advantage is that besides an overall score for providers, the method also provides a measure of how well each question performs in being able to distinguish between good and bad providers—how difficult the question is, how easy it is to guess the right answer, and how much difference there is between good and bad providers. In the first stage weights are generated for history and examination questions as well as the treatment assessments by the expert panel through the statistical technique itself. Once weights have been generated for every question, the provider is scored on a standardized distribution with mean zero and variance one.

Figure 1 shows the overall distribution of our competence measure in the entire sample along with the estimated standard error. The figure overlays the histogram of quality (on the left axis) and the standard error of our estimates. The line in the middle is our estimated quality (plotted against itself) while the curves on the extremes represent

Figure 1: Distribution of Competence and Standard Errors



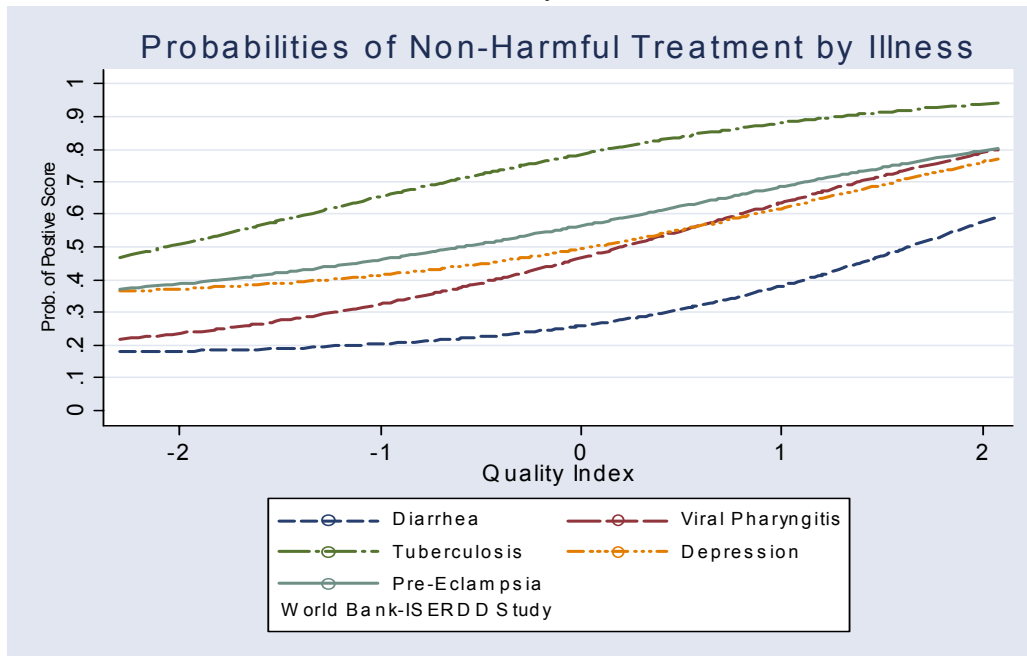
Note: This figure shows the distribution of competence in the entire sample with the estimated standard errors. Competence is measured on the horizontal axis (labeled Latent Variable (ML Estimate)) since this variable is a Maximum Likelihood estimate of a latent variable). The left vertical axis shows the density corresponding to the histogram in the diagram. The *right* vertical axis shows competence and confidence interval bands at the 95 percent confidence level. The line in the middle is competence plotted against itself, while the bounds on the outside correspond to the upper and lower confidence intervals, respectively. Techniques and estimation details are in Das and Hammer (2003b).

95 percent confidence intervals. One notable feature is that we have a much easier time distinguishing among good doctors (the confidence interval around our estimate is narrow) than among bad doctors (where the interval is much wider). This is in spite of the fact that we purposely included cases with very easy answers. Virtually no one, for example, asked to check the respiratory rate for patients complaining of coughing (in the two cases—viral pharyngitis and TB). In subsequent applications, we intend to add some prompted questions to get more information on the lower tail of the distribution.

III.3. Assessing Competence: Benchmarking

What does our measure of competence mean in terms of real-life characteristics of care? Figure 2 shows the probability of receiving care that is *not actually harmful* as a function of our measured index of competence for each of the five cases. For an infant with diarrhea, the probability that a provider who is highly competent in our ranking provides ORT and refrains from giving antibiotics or an anticholinergic is less than 60 percent and this drops to less than 30 percent for a provider of average competence. For viral pharyngitis, where the proper action is to do nothing, we find that 80 percent of

Figure 2: Probabilities of Non-Harmful Treatment by Illness



Note: This figure shows how the probability of a non-harmful treatment varies with competence. A “non-harmful” treatment is one that was graded positively by our panel of experts—the grading scheme was standardized so that a positive score referred to a treatment whereby the patient would be made “better-off” by visiting the provider. Each curve is based on a locally weighted least-squares estimation, which does not impose any parametric restriction on the form. The figure shows, for instance, that the probabilities of non-harmful treatment in TB (which corresponds either to referral or to treatment with a standard TB kit) increase from 50 to 95 percent moving from the least to the most competent provider.

highly competent providers would refrain from providing antibiotics and this decreases to less than 40 percent for the lowest ranks.

While poor treatment in these cases arises due to Type I errors (doing something when nothing should be done), performance with regard to Type II errors is also poor. In the case of tuberculosis, extremely competent providers are able to diagnose and treat appropriately 90 percent of the time, but this drops to 70 percent for the average and to less than 50 percent for the least competent. Perhaps most troubling is the probability of diagnosing and treating pre-eclampsia. Even among the most competent providers, this never exceeds 80 percent and drops to 55 percent for the average and to less than 40 percent for the least competent. That is, if the patient relied on advice from a single provider, extremely harmful consequences would result with a 20 percent probability, *even if she visited the most competent providers in our sample.*

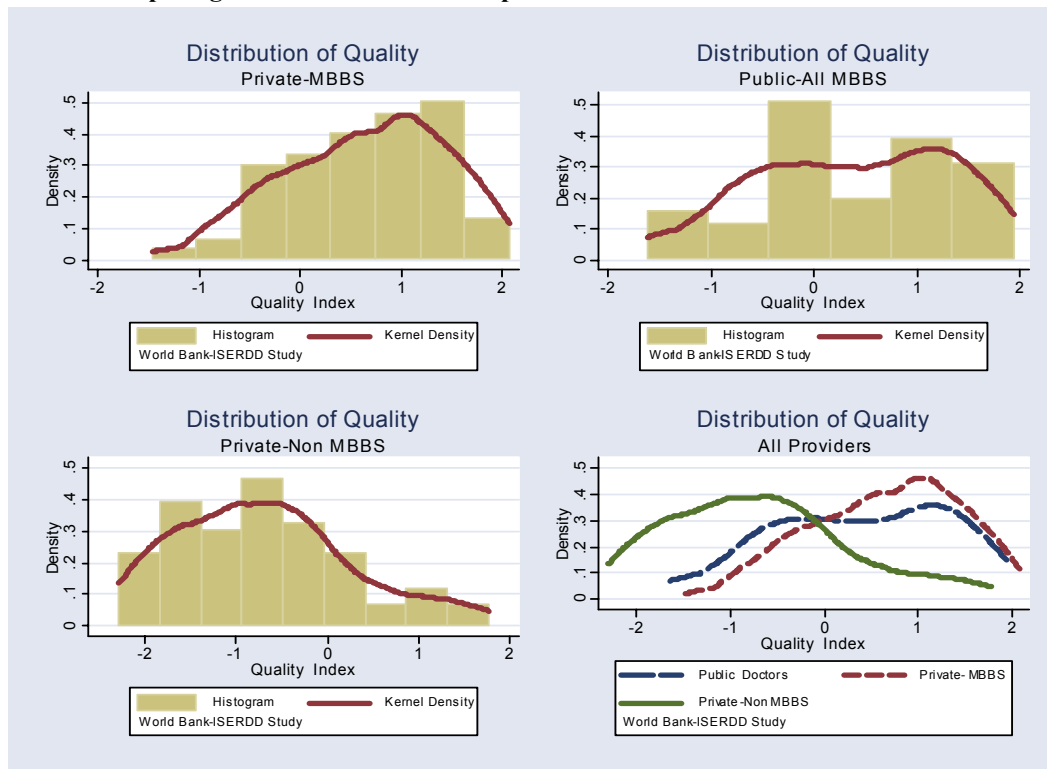
Our discussion of competence has so far concentrated on its measurement and the implications of our ranking scale for treatment. We next turn to a discussion of variation—how are providers distributed across rich and poor localities? Are there systematic differences between public and private providers? What are the typical options that a poor person faces in visiting a medical care provider?

III.4. What Does the Distribution of Competence Imply for Care?

The distribution of quality among doctors can be analyzed in several ways. One aspect of particular interest is the difference between the public and private sector. The private sector accounts for the vast majority of treatments in India as a whole and in Delhi in particular. This is well-known from the results of the 52nd round of the NSS and is reconfirmed in the ISERDD study of households (Das and Sanchez 2003). Perceptions of the relative quality of the two sectors vary. Some believe that health services in the private sector are better than services in the public sector, which is why the proportions going to a costly private facility are higher than those going to a free public facility, even among the poor. Others point out the high proportion of unqualified (or very minimally qualified) providers or “quacks” in the private sector that are much worse than staff in public facilities. As it turns out, everyone’s prejudices are confirmed by the data.

Figure 3 decomposes figure 1 into three types of providers: (1) private MBBS degree holders, (2) private non-MBBS providers, and (3) public doctors—all of whom have an MBBS. For each graph, we overlay the kernel density of competence on the underlying histogram. The results are striking. Private doctors form two skewed distributions—MBBS doctors displaying generally high competence (with a tail off to the left) and non-MBBS doctors displaying generally low competence (with a tail off to the right). Public doctors seem to be drawn from a bi-modal distribution. One group, which

Figure 3: Decomposing the Distribution of Competence



Note: This figure decomposes the distribution of competence corresponding to Figure 1 into private-MBBS, public (all MBBS) and private non-MBBS providers. The horizontal axis in all four figures corresponds to competence, and a kernel density plot is overlaid on the histogram of competence. We find that private MBBS providers are, on average, more competent than private non-MBBS providers, with public providers distributed evenly across the entire range, but bounded above by MBBS and below by non-MBBS providers in the private sector. Notably, there is also tremendous variation within qualifications—there are private MBBS providers who are worse than private non-MBBS providers, and public providers who are both better than private MBBS and worse than private non-MBBS providers.

happens to coincide largely with doctors in small clinics throughout the city, are somewhat better than their non-MBBS counterparts—but not by much. The other group, which generally come from larger public hospitals, are much better than either of these two groups but somewhat worse than their private MBBS counterparts—but again, not by much.

The options of patients who are poor are worse in a number of ways (Table 2). First, on average, poorer areas have a far higher percentage of private non-MBBS doctors (79 percent for low income compared to 31 percent for the rich) who have much lower average competence than their public or MBBS counterparts. Second, even within provider categories, less competent providers choose to locate in lower-income areas. Thus, if we restrict attention to MBBS doctors only, we find that average competence in low-income areas is almost one-half standard deviation lower than in rich areas. Exactly

Table 2: Distribution of Providers					
Provider category		Low income	Middle income	High income	Government hospitals
Non-MBBS private	Average competence	−0.86	−0.61	−0.32	
	% Total providers	79%	50%	31%	
MBBS-private	Average competence	0.46	0.27	0.92	
	% Total providers	18%	44%	61%	
Public (MBBS)	Average competence	−0.57	0.62	0.28	0.58
	% Total providers	3%	6%	5%	100%
<p><i>Note:</i> This table shows the distribution of competence and qualifications across areas with different average incomes. The incomes were computed from a consumption survey administered to households in the ISERDD study where households were picked randomly from every community. Of the seven communities, three are low income, two are middle income, and two are high income. Large government hospitals are categorized in a separate category rather than assigned to any income group. Competence refers to the IRT measure and the percentage of total providers is computed from a census of all providers within a 15-minute walking radius of the households in the ISERDD study. Non-MBBS private providers include (but are not restricted to) Bachelor of Ayurvedic Medicine and Surgery (BAMS), Bachelor of Integrated Medicine and Surgery (BIMS), Bachelor of Unani Medicine and Surgery (BUMS), Registered Medical Practitioners (RMP) and Practicing Medical Practitioners (PMP). For the purpose of the study, a provider was defined as “any individual or institution providing medical care on a systematic basis.”</p>					

the same results hold for non-MBBS providers, to the extent that such providers locate in richer areas, they tend to be far more competent.

While average provider competence rises with qualifications and better neighborhoods, there is tremendous variation *within* these categories. A simple ANOVA decomposition reveals that less than one-half of the variation (43 percent) in provider competence is accounted for by differences across neighborhoods, qualifications, and institutional affiliation (public or private), with the remainder accounted for by intra-locality and intra-qualification variation. Consequently the problem of low *average* competence in poor localities is compounded by high *variation* in competence along observed dimensions (locality, qualifications, and whether the doctor is in the public sector) by the individual—a large proportion of competence is idiosyncratic so that there is no guarantee that a MBBS doctor in a rich neighborhood is, in fact, highly competent.

Given the low average competence and high variation in competence of private providers in poor neighborhoods, public doctors might be an attractive option. Unfortunately, this is not borne out by the data. For a poor individual, the differences between the public and private sector do not mean much. In poor neighborhoods, the average competence of public doctors is extremely low—this increases with income, but

is comparable to that of private MBBS doctors only for government hospitals. One option is to visit only government hospitals in affluent areas, which are theoretically available to anyone. In terms of observed behavior, however, low-income families rarely use these hospitals except in an emergency. Long distances, high travel costs, long waiting times and discourtesy inhibit the use of these facilities on a regular basis by the poor.

Where the two sectors differ, at least in terms of what they *say* they would do during the vignettes, is in the kind of treatment they provide. For diarrhea, the chances of being successfully treated (using ORT) is entirely determined by the quality of doctor and is unrelated to either their location or whether they are in the public or private sector. However, in poor areas, the vignettes indicated that good treatment is often accompanied by unnecessary drugs in the private sector, whereas in richer neighborhoods this is not the case. Similar results hold for viral pharyngitis, where providers in the public sector and rich neighborhoods are less likely (by 30 percent and 25 percent, respectively) to provide antibiotics compared to those operating privately in poor localities. Hence, in cases where the appropriate response is to do nothing or to refer, public sector providers are much more likely to respond appropriately.⁸ There may be a general tendency to do less in the public sector and when that is the right thing to do, they perform better.

These results, however, are based on our measure of competence. A natural question to ask from this analysis is how competence translates into practice. Does the extent of Type I and Type II errors hold in the clinic? We turn to this question using data from direct clinical observation.

IV. What Providers Do: The Relationship between Competence and Practice

The second type of data was collected by direct observation during providers' clinical practice. Although observers were not medical professionals, and this limited our ability to judge details of the treatments, it also avoided ethical problems were a doctor to observe improper treatment. For every provider-patient interaction the observer filled in a simple form that included the number of questions the provider asked (with more details for cases of diarrhea and short-duration cough), and noted whether a physical examination took place and, if so, which examination was performed, how much time was spent with the patient, which drugs were prescribed (to be evaluated by our panel of physicians in South Asia and the U.S. afterwards), and recorded the fee charged. Of a total of 4,100 interactions observed, 1,770 observations were in public sector facilities.

What does the average patient look like and what does the average provider do? The magic number appears to be three: The median (mean) patient arrives with 3 (8) days

⁸ Details can be found in Das and Hammer (2003a).

of illness behind him, presenting with pain, fever or cough (or some combination). In the median interaction the provider then asks 3 (3.2) questions regarding the illness and performs some examinations (which would probably involve using a stethoscope and checking the patient's temperature). The patient is then provided with 3 (2.6) different medicines (providers dispense rather than prescribe medications in 69 percent of all interactions) and the interaction is over in 3 (3.8) minutes. Patients are seldom referred (less than 7 percent), given instructions (50 percent of the time), or offered guidance regarding follow-up (35 percent of the time).

In a sense the results below on the differences between public and private providers represent a second-order problem. Clearly a key issue is the paucity of care in terms of careful diagnosis and treatment in the median interaction irrespective of whether the care was received in the public or private sector. Given that the average case in the vignettes took 15 minutes to complete (without the provider actually performing an examination) the sharply reduced consultation time is itself an indication that there are wide differences between competence and practice.

IV.1. Competence and Practice: Is There a Difference?

First, it is clear that our measure of competence *does* affect practice with more competent providers performing better along a number of dimensions (Table 3). Thus, with higher competence more questions are asked, more examinations are performed, more time is spent, and more money is charged. These increases, though, are small; moving from the lowest to the highest quintile of competence increases the number of history questions by 1.8, the probability of an examination by 18 percent, and the time spent by 1 minute. Competence is thus correlated with practice. What then is the extent of the wedge?

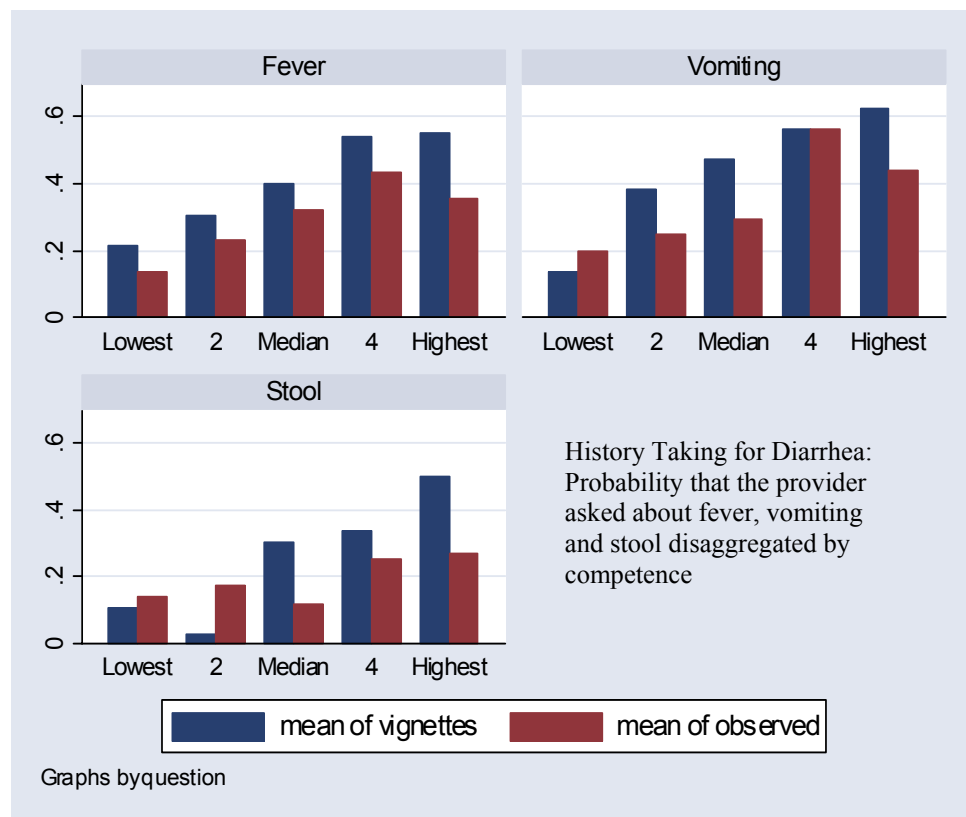
To evaluate this gap, we compare history taking for two illnesses, both represented in the vignettes and direct clinical observation. For cases of diarrhea and cough, the observer was asked to specifically check if the provider asked about fever, vomiting and stool (in the case of diarrhea) and fever, chest pain, and expectoration (in the case of cough). Since these questions were graded in the vignettes, a direct comparison is possible, shown in figure 4.⁹

⁹ A similar comparison for viral pharyngitis can be found in Das and Hammer (2003b).

Table 3: Practice and Competence

Quintiles of ability	Total history questions	Probability of examination	Total time spent (minutes)	Fees charged (Rs.)
Lowest	2.59	58.3	3.47	15.2
2 nd Group	2.94	62.5	4.04	24.3
Median	2.75	56.2	2.97	13.5
4 th Group	3.62	65.0	4.43	24.1
Highest	4.35	76.7	4.49	25.2

Note: This table disaggregates provider practice (observed in the clinic) by competence measured in the vignettes. History questions refer to the number of questions regarding the illness that the provider asked the patient. An examination consists of any physical contact between the provider and the patient or the use of measuring instruments, such as a thermometer, sphygmomanometer, or stethoscope. Note that an examination only implies that the device was used, not that the device was used correctly. Fees charged refers to the total payment at the end of the interaction. Since this sample includes public providers, the increase with competence is much less than for private providers only.

Figure 4: Competence and Practice in History Taking (Diarrhea)

Note: This figure shows the probability of asking a question in the vignettes and actual observation. Three questions were both asked in the vignettes and specifically noted in observation—whether the child had a fever (top left), whether the child had experienced any vomiting (top right) and what the nature of the stool was (bottom left). The horizontal axis shows quintiles of competence and the vertical axis the probability that the question was asked in the vignettes (first bar) and in observation (second bar). The probability of asking a question increases in competence for both vignettes and observation. However, this probability is strictly lower in observation, and the gap widens with competence.

Two observations stand out. First, fewer providers ask patients about these symptoms in observed interactions compared to vignettes. This decline is particularly noticeable for all the questions related to diarrhea and for the question related to fever in the case of viral pharyngitis (not shown). For instance, 50 percent of all providers ask about fever in diarrhea in the vignettes and this drops to 28 percent in observation. For viral pharyngitis, there is a dramatic drop in questioning about fever from 58 percent (vignettes) to 16 percent (observation). Second, the impact of competence on the probability of asking questions is much lower in observation than in the vignettes—in other words, the gap between competence and practice *widens* as one moves up the scale. For viral pharyngitis, a one-standard deviation increase in competence implies a 26 percent increase in the probability of enquiries concerning fever during vignettes, but only a 4 percent increase in observation.

The same results hold for treatment patterns.¹⁰ In the case of viral self-limiting illnesses, providers *know* that they should not treat with anti-infectives (vignettes) but in treatment, the probability of treating with such drugs increases by close to 20 percent for the entire sample, both in the case of diarrhea and coughs with less than a three-day history. Moreover, as in our results regarding history taking, the gap between competence and practice is wider for more competent providers.

These results are troubling. It has long been thought that the primary problem with health care in low-income countries is related to the paucity of care (see, for instance, Box 8.10 in World Bank, 2003). Yet here we have a puzzle—despite the fact that the population is well-served by providers (with more than 70 to choose from within a 15-minute walk from every household) the quality of the provider-patient interaction is poor, with less than 3 minutes spent and 3 questions asked on average. Two related factors increase the severity of the problem. First, the effect of training and competence on practice is severely attenuated, despite higher competence levels, and the behavior of “better” providers is not much better than that of their “worse” counterparts. Second, private practice does not help, at least as much as we would like it to. Even strong incentives to providers (by linking profits to the number of patients seen) does not lead to as high an improvement in quality of care as one would have expected given the high level of competition. Although more competent providers do provide better care in terms of process indicators such as time and history-taking, this increase is small and there is a wide gap between what they know they should do and what they actually do, and this gap actually widens with competence.

¹⁰ Das and Hammer (2003b).

To improve performance we need a better understanding of factors that affect this gap. In terms of policy, this is akin to saying that we do not need to increase the number of providers, or provide them with better training, but we do need to make sure that they practice in a manner that is consistent with their competence. On the flip side, the wedge also implies that increasing the number of providers in an area is not enough and neither is increasing the level of training *unless* the quality of health care actually provided in practice also increases. Whether or not this is the case depends, of course, on the manner in which providers use the knowledge they possess.

In the final section, we provide preliminary evidence that this wedge responds to incentives. Part of the gap between competence and practice is related to the specific incentives that providers face. In the public sector, with salaried jobs and little monitoring, the incentive to provide high-quality care is sharply reduced. In the private sector, there are very powerful incentives to provide *one particular* type of care—care that is demanded by the patient and that will cause him/her to return to the provider, whether or not this course of treatment is justified by standard-of-care protocols.

IV.2. Incentives and the “Wedge”: Evidence from Public/Private Providers

We start by examining basic process indicators—the time spent with patients and the probability of an examination. Table 4 presents the same information as Table 3 but is broken down into public and private practices. Note that at every quality level public providers perform worse than private—they ask less history questions, are less likely to perform a physical examination, and spend less time with the patient. The gradient (the rate of increase up the quality ladder) is higher in the public sector than in the private so that there is some convergence at the top end. However, a low-quality public provider—the kind a poor person is likely to see (as indicated in Table 1) is prone to ask fewer than two questions, spend less than two minutes with each client, and does a physical exam of any kind in less than one-third of the visits. Perhaps it is not so surprising that even poor people prefer to pay the Rs. 20 or so for a low-quality provider rather than go to a public clinic with a qualified provider (and this is just in terms of clinical care without considering any possible differences in the degree of courtesy shown).

Some simple explanations for the differences might be that public providers see patients with a different set of conditions or generally less severe cases of the same conditions. This would imply that the quality of care is similar since less complicated cases might take less time and effort to resolve. Neither explanation is correct. The distribution of symptoms of the patients in both sectors was remarkably similar. Approximately 35 to 40 percent were for “pain” in both the public and private sector, and about the same for fever and/or cough. Another eight less frequent illness categories are

Table 4: Practice and Competence across Sectors					
Quintiles of ability	Type of provider	Total history questions	Probability of examination (percent)	Time spent (minutes)	Fees charged (Rs.)
Lowest	Private	2.93	70	4.22	21.5
	Public	1.72	28	1.56	0
2 nd Group	Private	3.37	71	4.76	34.6
	Public	1.88	41	2.28	0
Median	Private	3.55	75	4.06	32.8
	Public	2.17	42	2.19	0
4 th Group	Private	3.67	81	5.39	44.0
	Public	3.55	41	3.10	0
Highest	Private	4.71	81	5.89	57.3
	Public	4.04	72	3.26	.01
<p><i>Note:</i> This table disaggregates provider practice (observed in the clinic) by competence measured in the vignettes and sector (public/private). History questions refer to the number of questions regarding the illness that the provider asked the patient. An examination consists of any physical contact between the provider and the patient or the use of measuring instruments, such as a thermometer, sphygmomanometer, or stethoscope. Note that an examination only implies that the device was used, not that it was used correctly. Fees charged refers to the total payment at the end of the interaction. Finally, public providers are those who were observed in their public practice and need not be providers who work only in the public sector.</p>					

almost identical between the sectors. Also, as measured by the number of days that people were sick before they sought medical treatment, the public sector tended to see more severe cases of the same conditions—people waited longer to go to public facilities (the median number of days a person was sick prior to the interaction in the public sector is seven compared to three for the private sector). Consequently, the case-mix bias should operate in the opposite direction to what we see in the data, and what we actually observe is probably a *lower bound* of the difference in process indicators between the public and private sector.

The same point can be made more rigorously in a regression context. Table 5 shows a set of regression results in which the time spent on patients and the probability of doing an examination are both explained by competence, public or private practice, type of neighborhood the practice is in, and whether, if public, the practice is in the out-patient ward of a general hospital. In all regressions we include a full set of dummies for the type of illness that the patient reported with, to control for potential selection issues.

Without controlling for the possibility that public providers have a much larger workload that would force them to spend less time with their patients regardless of their inclination, we find that they do, in fact, spend considerably less time with their patients (Column 1). Given that the average amount of time spent in the entire sample is only about 3.7 minutes the fact that the difference between public and private providers in

clinics is also 3 minutes means that in a poor area, say, a private provider would average 5 minutes per client, the public provider only 2 minutes. Providers in public hospitals do spend about a minute more per client than other public providers located in smaller government dispensaries throughout the city, but even this is significantly less than the private sector.

When we control for the total number of patients seen per hour (during the period of observation) in Column 2, a substantial fraction of this effect is reduced. A heavier workload does explain some of the difference, but even with this accounted for, public providers still see patients for a significantly shorter period of time.¹¹ Columns 3 and 4 repeat this analysis for the probability of conducting a physical exam. Again, public providers are less likely to conduct any examination than their private counterparts. The difference remains large and significant once even controlling for the case load *and* the time spent with the patient, although it decreases from a 40 percent differential to 28 percent. Thus, even if public providers spent as much time with their patients as their private counterparts, they would still be less likely to undertake any sort of physical examination. Given that they are likely to see sicker patients, this is even more striking.

To analyze in greater detail the impact of poor incentives for the provision of health care, we can follow a similar strategy to our discussion on the gap between competence and practice above. Specifically, if we select only cases of diarrhea and cough with a short history (less than 3 days) we can directly compare the probability of asking specific questions in the vignettes to performance in practice. The results from this exercise are striking.¹² Again, the relationship between competence and the probability of asking relevant questions is much lower in practice than in the vignettes. The results are particularly dramatic for vomiting and questions about stool in diarrhea and fever and expectoration for cough, with the number of questions asked in practice by highly competent providers being less than one-half of those in vignettes.

Next, in addition to the practice penalty we also observe a public-sector penalty—the likelihood that a provider in the public sector uses his/her full knowledge in practice is much lower compared to providers in the private sector. For viral pharyngitis, the penalty is severe. In the vignettes we do not observe a significant difference between public and private providers in their probability of asking these questions. In practice however, the penalty on public providers is always significant and ranges from a 6.6

¹¹ Further, it is not entirely clear that the case load should actually be controlled for. If, for instance, the provider is absent from work, the case load will be higher, but this is an endogenous choice that should not be accounted for. In this sense, the estimate from Column 1 is an upper bound of the difference and that of Column 2 is the corresponding lower bound.

¹² See Das and Hammer (2003b) for details.

Table 5: Practice and Competence (Regression Results)				
	(1)	(2)	(3)	(4)
	Time spent with patient (not including case load)	Time spent with patient (including case load)	Marginal probabilities of examination (not including case load or time spent)	Marginal probabilities of examination (including case load & time spent)
Competence	0.472 [0.224]**	0.569 [0.201]***	0.053 [0.021]**	0.058 [0.023]**
Public provider	-3.015 [0.379]***	-1.620 [0.478]***	-0.416 [0.061]***	-0.274 [0.058]***
Interaction: Public x competence	0.115 [0.278]	-0.682 [0.372]*	0.056 [0.042]	0.003 [0.044]
Middle-income neighborhood	-0.899 [0.443]**	-1.387 [0.416]***	0.101 [0.052]*	0.062 [0.051]
High-income neighborhood	0.264 [0.614]	-0.318 [0.551]	0.040 [0.061]	-0.007 [0.053]
Government hospitals	0.935 [0.440]**	0.280 [0.441]	0.261 [0.070]***	0.204 [0.066]***
Number of patients per hour		-0.120 [0.033]***		-0.012 [0.004]***
Total time spent with patient: minutes				0.012 [0.004]***
Controls for type of illness	Yes	Yes	Yes	Yes
Constant	4.955 [0.408]***	6.082 [0.460]***		
Observations (number of interactions)	4077	3998	4092	3998
R-squared	0.16	0.21		
<p><i>Note:</i> These regression results show the relationship between two process indicators (time spent with the patient and the probability of an examination) and competence, income, and sector. Columns (1) and (2) are ordinary least squares regressions and columns (3) and (4) are probit regressions. Standard errors are reported in [brackets], (***) denotes significance at the 1 percent level, (**) at the 5 percent level and (*) at the 1 percent level. Errors are clustered at the level of the provider, since there are multiple interactions for every provider. All regressions control for the symptoms that the patient reported with. These include fever, pain, cold, cough, diarrhea, vomiting, injury, weakness, dermatological problems, and mental health issues.</p>				

percent decline (expectoration) to a 12 percent decline (fever). The results for diarrhea are less emphatic, although we do observe a 50 percent difference in questions about fever similar to the case above.¹³

¹³ This is due to the presence of a penalty in the vignettes as well. In the vignettes providers in the public sector often said that they would refer the case to the pediatric ward and thus did not ask any questions at all, although this was never observed in practice. Restricting the sample to those who did not refer gives us the same pattern of results as viral pharyngitis.

This evidence shows that public providers do much less than private providers, even when controlling for workload and when two important features of case load either do not differ (types of problems) or would tend to lead to more time and examinations being necessary (severity). In the case of diarrhea the decrease in the probability of asking about fever and the nature of the stool implies that a public provider would probably be unable to differentially diagnose dysentery from viral diarrhea, with potentially life threatening consequences. The errors thus committed in the public sector are Type II errors (not doing something when it should be done), reflecting presumably differences in incentives between providers in the two sectors.

The evidence from the vignettes seems to indicate that private providers, while doing more for their patients, are doing too much (Type I errors). In an analysis similar to that above we find that after controlling for qualifications and the type of illness, public providers prescribe 0.13 less antibiotics (this amounts to 0.2 standard deviations of the distribution of antibiotics prescribed) and 0.53 less drugs overall (amounting to almost 0.4 standard deviations of the distribution of the number of drugs prescribed).¹⁴

Proceeding to the diarrhea and viral pharyngitis cases our results are less clear. We *do* find clear evidence that private providers prescribe more antibiotics in both of these cases compared to their performance in the vignettes; in fact, competence does not affect the probability of receiving antibiotics in practice, although it strongly decreases the probability of receiving such drugs in the vignettes. In the case of diarrhea the difference between private and public providers is **not** significant in the provision of antibiotics and **is** significant at the 10 percent level of confidence for the provision of anti-diarrheals. For coughs and fevers of short duration (less than 3 days) we find that public providers prescribe fewer antibiotics, but only at higher levels of competence.

These results suggest that the gap between public and private providers is much higher in observation than in vignettes. However, the gap is in opposite directions for effort in treatment and appropriateness of treatment. In terms of effort exerted providers in the public sector perform worse. In contrast, when we look at treatment patterns, the private sector has a tendency to over-prescribe and medicate for disorders that are self-limiting. In other words, the public sector is more prone to Type II errors and the private sector to Type I errors. There is hence a further dimension to the standard public-private comparison, whereby the private sector is characterized as high quality-high cost and the public as low quality-low cost, and this relates to the multi-dimensional nature of quality. If we view quality solely as effort induced, the private sector clearly elicits higher quality from its providers. Alternatively, if we view quality as the extent to which the provider

¹⁴ The interested reader is referred to Das and Hammer (2003b).

behaves in the patient's best interest, the comparison is not as clear-cut as before—while the private sector does more, it may be doing too much.

V. Discussion and Conclusion

This analysis has important implications for health policy in India. But it also points toward the need for radical rethinking of health care structures for the poor. An important idea to emerge from the influential Alma-Ata declaration (World Health Organization 1978) was that primary health care outcomes would improve with increased access. Thus, the declaration states that:

“Primary health care relies, at local and referral levels, on health workers, including physicians, nurses, midwives, auxiliaries and community workers as applicable, as well as traditional practitioners as needed, suitably trained socially and technically to work as a health team and to respond to the expressed health needs of the community.” (Declaration of Alma Ata, World Health Organization, 1978)

Consistent with this declaration, Indian health policy sought to actively improve access to health care (especially in rural areas) through a two-pronged strategy. First, a tier of health workers was created through rapid training programs (Registered Medical Practitioners), who would ostensibly act at the front line of a triaging system. Second, it was argued that by mainstreaming providers practicing traditional medicine (thus creating the categories of BAMS, BIMS and BUMS), cheap and efficient health care could be provided for the poor. Our analysis suggests that the key assumptions on which this policy was based—that providers would remain in the areas they were trained and would practice in a manner consistent with their training—were overly optimistic.

On the one hand, trained providers migrated from their initial center of training to Delhi. Thus, in our census of providers we find that among RMPs and providers with informal degrees, 80 percent were trained in states other than Delhi and this drops to 65 percent for MBBS. Competence among providers from out of state is 0.2 standard deviations less than those trained in Delhi, reflecting their envisioned position in the health triage system.

On the other hand, the specific types of practices followed in the clinic quickly deviated from the standard protocol of a triaging system. Thus, both in the vignettes and in clinical observation, providers who were supposed to act as the first step of the triage system failed to perform either of their two functions—they were not able to diagnose and treat minor illnesses effectively, and they were not able to diagnose and refer complicated illnesses. The predominant mode of practice was treatment in all cases, with an average of three medicines provided in every encounter. More troublingly, the referral system worked in exactly the opposite manner to that envisioned—overall referrals were

low (less than 7 percent) and the probability of a referral *increased* with the competence of a provider. Providers with lower levels of competence were more likely to treat rather than refer illnesses. Further, providers trained in alternative systems of medicine were unable to insulate themselves from the exceptional penetration of the pharmaceutical markets in Delhi. We find, for instance, that although 28 percent of our sample of providers hold degrees in alternative medicines, less than 5 percent of all medicines used were Ayurvedic or Unani or Herbal remedies. Modes of practice therefore evolved in a manner that was inconsistent with the original ideas presented in the Alma Ata declaration.

Our findings show that the effect of creating a system with some highly qualified and some less qualified providers is even more pernicious due to the adverse impact on the practice of competent providers. What the data suggest (recall the widening of the gap between competence and practice at higher levels of competence) is that in the health market greater (and more varied) competition leads to a *worsening* of treatment patterns as more competent providers are faced with the choice of either catering to their patient's expectations or losing profits.

This sets up an interesting contrast between the incentives in the different institutional settings for medical care. Providers on salary have little incentive to spend much time on each patient and their practice confirms this. Providers paid on a fee-for-service basis have a strong incentive to cater to the particular demands of the client and will resort to treatment protocols that are not medically sound. Also, if the provider sells medicines there is an incentive to prescribe more than is necessary (85 percent of all interactions in the private sector resulted in some medicine being dispensed as opposed to prescribed). The evidence in our study indicates that both are prone to errors but of different kinds. Public providers exert less effort and treat symptomatically, if at all. Private providers are prone to over-medication. It is unclear what can be done to change these incentives in a systematic way.

If patients know themselves to be ill in such a way that primary care is needed, they are better off visiting a private provider. If they are unsure of the severity of their illness, or alternatively, are aware that their illness is severe and requires specialist attention, the public sector is the better choice. Unfortunately, most patients are unaware of the nature of their problems prior to a diagnosis. One strategy is to try to improve the level of medical education among patients. Since the large majority of care is in the private sector and since that sector is prone to follow the demands of the patient, a better informed patient is an obvious asset. This is better for the consumers themselves but also, inasmuch as they will use fewer antibiotics, there is a direct externality in the form of slower development of drug resistance. Less is known about the specific means that can be

adopted in bringing about this change. There is some evidence that demonstration effects (through high standard-of-care facilities) change patient expectations, and if this is true, then using provider-patient interactions in public hospitals to demonstrate standard treatment protocols may prove effective.

The behavior of medical care providers is a seriously understudied area of research, particularly in the economics literature in developing countries. Since much of this behavior is opportunistic, in the sense that it reflects the underlying structure of incentives, economic insights are required to better structure analysis in this sector. While our results have been largely negative, suggesting that there are no “quick fixes,” this paper and the research it draws on represent the first attempts to address this problem in a more systematic manner.

Bibliography

The word “processed” describes informally reproduced works that may not be commonly available through library systems.

Das, Jishnu, and Jeffrey Hammer. 2003a. “Which Doctor: Combining Vignettes and Item Response to Measure Doctor Quality.” Development Research Group, World Bank, Washington, D.C. Processed.

Das, Jishnu, and Jeffrey Hammer. 2003b. “Money for Nothing: The Dire Straits of Medical Practice in India.” Development Research Group, World Bank, Washington, D.C. Processed.

Das, Jishnu, and Carolina Sánchez-Páramo. 2003. “Short but not Sweet: New Evidence on Short Duration Morbidities from India.” Policy Research Working Paper 2971. World Bank, Development Research Group, Washington, D.C.

Das, Veena, and Ranendra K. Das. 2003. “Pharmaceuticals in urban ecologies: The register of the local.” Processed.

Institute of Socio-Economic Research on Development and Democracy, Various years. Occasional Research Paper Series. Delhi.

Collier, Paul, Stefan Dercon, and John Mackinnon. 2003. “Density versus Quality in Health Care Provision: Using Household Data to Make Budgetary Choices in Ethiopia.” *The World Bank Economic Review* 16(3): 425–48.

Greenhalgh, Trisha. 1987. “Drug prescription and self-medication in India: an exploratory survey.” *Social Science and Medicine* Vol. 25: 307–18.

Kamat, Vinay R. and Mark Nichter. 1998. “Pharmacies, self-medication and pharmaceutical marketing in Bombay, India.” *Social Science and Medicine* 47: 779–94.

Lavy, Victor, and Jean-Marc Germain. 1994. “Quality and Cost in Health Care Choice in Developing Countries.” Working Paper 105. Living Standards Measurement Study, World Bank, Washington, D.C.

- Leonard, Kenneth L. 2003. "African Traditional Healers and Outcome-Contingent Contracts in Health Care." *Journal of Development Economics* 71(1).
- Leonard, Kenneth L., and Melkiory C. Masatu. 2003. "Comparing Vignettes and Direct Clinical Observation in a Developing Country Context." University of Maryland, College Park, Maryland. Processed.
- Phadke, Anant. 1998. *Drug Supply and Use: Toward a Rational Policy in India*. New Delhi: Sage Publications.
- Rethans, J. J., F. Sturmans, R. Drop, C. P. M. van der Vleuten, and P. Hobus. 1991. "Does competence of general practitioners predict their performance? Comparison between examination setting and actual practice." *British Medical Journal* 303 (6814): 1377–80.
- World Bank. 2003. *World Development Report 2004: Making Service Work for Poor People*. Washington, D.C.: World Bank.
- World Health Organization. 1978. "The Alma-Ata Declaration." Geneva: World Health Organization.